



SEQUENCE LISTING

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to Commissioner of Patents and Trademarks, Washington, D.C. 20231, on November 9, 1992.

Ellen J. Kapinos

(1) GENERAL INFORMATION:

(i) APPLICANT: Israel, David
& Wolfman, Neil M.

(ii) TITLE OF INVENTION: Recombinant Bone Morphogenetic Protein Heterodimers, Compositions and Methods of Use.

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
(B) STREET: 87 Cambridge Park Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02140-2387

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Tape
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Kapinos, Ellen J.
(B) REGISTRATION NUMBER: 32,245
(C) REFERENCE/DOCKET NUMBER: GI-5192B

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 617-876-1170
(B) TELEFAX: 617-876-5851

!) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1607 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 356..1543

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

(A) LENGTH: 396 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val
1 5 10 15

Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys
20 25 30

Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu
35 40 45

Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys
50 55 60

Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu
65 70 75 80

Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp
85 90 95

His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe
100 105 110

His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr
115 120 125

Thr Arg Arg Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe
130 135 140

Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala
145 150 155 160

Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile
165 170 175

Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu
180 185 190

Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Thr Phe Asp
195 200 205

Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His
210 215 220

Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser
225 230 235 240

Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser
245 250 255

Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys
260 265 270

Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His Lys Gln

GTCGACTCTA GAGTGTGTGT CAGCACTTGG CTGGGGACTT CTTGAACCTG CAGGGAGAAT	60
AACTTGCAGCA CCCCCACTTG CGCCGGTGCC TTTGCCCGAG CGGAGCCTGC TTGCCCCATCT	120
CCGAGCCCCA CCGCCCCCTCC ACTCCTCGGC CTTGCCCGAC ACTGAGACGC TGTTCCCAGC	180
GTAAGAAAGAG AGACTGCGCG GCCGGCACCC GGGAGAAGGA GGAGGCAAAG AAAAGGAACG	240
GACATTCGGT CCTTGCGCCA GGTCCCTTGA CCAGAGTTT TCCATGTGGA CGCTCTTCA	300
ATGGACGTGT CCCCGCGTGC TTCTTAGACG GACTGCGGTG TCCTAAAGGT CGACC ATG	358
Met	
1	
GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC CTC	406
Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Pro Gln Val Leu	
5 10 15	
CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC	454
Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys Phe	
20 25 30	
GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC	502
Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu Val	
35 40 45	
CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG	550
Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys Gln	
50 55 60 65	
AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC	598
Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu Asp	
70 75 80	
CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC CAC	646
Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp His	
85 90 95	
CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC CAC	694
Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe His	
100 105 110	
CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA ACC	742
His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr Thr	
115 120 125	
CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT ATC	790
Arg Arg Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe Ile	
130 135 140 145	
ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT TTA	838
Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala Leu	
150 155 160	
GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC ATA	886
Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile	
165 170 175	
AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG GAC	934
Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu Asp	

180	185	190	
ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA ACT TTT GAT GTC Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Thr Phe Asp Val	195 200	205	982
ACC CCC GCT GTG ATG CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT GGA Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His Gly	210 215	220 225	1030
TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC AAG Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser Lys	230	235 240	1078
AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC TGG Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser Trp	245	250 255	1126
TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA GGG Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys Gly	260	265 270	1174
CAT CCT CTC CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC AAA CAG CGG His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys His Lys Gln Arg	275	280 285	1222
AAA CGC CTT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC TTC Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp Phe	290 295	300 305	1270
AGT GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His	310	315 320	1318
GCC TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu	325	330 335	1366
AAC TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn	340	345 350	1414
TCT AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile	355	360 365	1462
TCG ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA TTA AAG AAC TAT Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr	370	375 380 385	1510
CAG GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC TAGTACAGCA AAATTAAATA Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg	390	395	1563
CATAAAATATA TATATATATA TATATTTAG AAAAAAGAAA AAAA			1607

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

275	280	285	
Arg Lys Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp			
290	295	300	
Phe Ser Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr			
305	310	315	320
His Ala Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His			
325	330	335	
Leu Asn Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val			
340	345	350	
Asn Ser Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala			
355	360	365	
Ile Ser Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn			
370	375	380	
Tyr Gln Asp Met Val Val Glu Gly Cys Gly Cys Arg			
385	390	395	

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1954 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 403..1626

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTCTAGAGGG CAGAGGAGGA GGGAGGGAGG GAAGGAGCGC GGAGCCCCGGC CCGGAAGCTA	60
GGTGAGTGTG GCATCCGAGC TGAGGGACGC GAGCCTGAGA CGCCGCTGCT GCTCCGGCTG	120
AGTATCTAGC TTGTCTCCCC GATGGGATTC CCGTCCAAGC TATCTCGAGC CTGCAGCGCC	180
ACAGTCCCCG GCCCTCGCCC AGGTTCACTG CAACCGTTCA GAGGTCCCCA GGAGCTGCTG	240
CTGGCGAGCC CGCTACTGCA GGGACCTATG GAGCCATTCC GTAGTGCCAT CCCGAGCAAC	300
GCACTGCTGC AGCTTCCCTG AGCCTTCCA GCAAGTTGT TCAAGATTGG CTGTCAAGAA	360
TCATGGACTG TTATTATATG CCTTGTTC TGTCAAGACA CC ATG ATT CCT GGT	414
Met Ile Pro Gly	
1	
AAC CGA ATG CTG ATG GTC GTT TTA TTA TGC CAA GTC GTC CTG CTA GGA GGC	462
Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val Leu Leu Gly Gly	
5	
10	
15	
20	

GCG AGC CAT GCT AGT TTG ATA CCT GAG ACG GGG AAG AAA AAA GTC GCC Ala Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys Lys Lys Val Ala 25 30 35	510
GAG ATT CAG GGC CAC GCG GGA GGA CGC CGC TCA GGG CAG AGC CAT GAG Glu Ile Gln Gly His Ala Gly Gly Arg Arg Ser Gly Gln Ser His Glu 40 45 50	558
CTC CTG CGG GAC TTC GAG GCG ACA CTT CTG CAG ATG TTT GGG CTG CGC Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met Phe Gly Leu Arg 55 60 65	606
CGC CGC CCG CAG CCT AGC AAG AGT GCC GTC ATT CCG GAC TAC ATG CGG Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro Asp Tyr Met Arg 70 75 80	654
GAT CTT TAC CGG CTT CAG TCT GGG GAG GAG GAG GAA GAG CAG ATC CAC Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu Glu Gln Ile His 85 90 95 100	702
AGC ACT GGT CTT GAG TAT CCT GAG CGC CCG GCC AGC CGG GCC AAC ACC Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser Arg Ala Asn Thr 105 110 115	750
GTG AGG AGC TTC CAC CAC GAA GAA CAT CTG GAG AAC ATC CCA GGG ACC Val Arg Ser Phe His His Glu Glu His Leu Glu Asn Ile Pro Gly Thr 120 125 130	798
AGT GAA AAC TCT GCT TTT CGT TTC CTC TTT AAC CTC AGC AGC ATC CCT Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu Ser Ser Ile Pro 135 140 145	846
GAG AAC GAG GTG ATC TCC TCT GCA GAG CTT CGG CTC TTC CGG GAG CAG Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu Phe Arg Glu Gln 150 155 160	894
GTG GAC CAG GGC CCT GAT TGG GAA AGG GGC TTC CAC CGT ATA AAC ATT Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His Arg Ile Asn Ile 165 170 175 180	942
TAT GAG GTT ATG AAG CCC CCA GCA GAA GTG GTG CCT GGG CAC CTC ATC Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro Gly His Leu Ile 185 190 195	990
ACA CGA CTA CTG GAC ACG AGA CTG GTC CAC CAC AAT GTG ACA CGG TGG Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn Val Thr Arg Trp 200 205 210	1038
GAA ACT TTT GAT GTG AGC CCT GCG GTC CTT CGC TGG ACC CGG GAG AAG Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp Thr Arg Glu Lys 215 220 225	1086
CAG CCA AAC TAT GGG CTA GCC ATT GAG GTG ACT CAC CTC CAT CAG ACT Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His Leu His Gln Thr 230 235 240	1134
CGG ACC CAC CAG GGC CAG CAT GTC AGG ATT AGC CGA TCG TTA CCT CAA Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg Ser Leu Pro Gln 245 250 255 260	1182

GGG AGT GGG AAT TGG GCC CAG CTC CGG CCC CTC CTG GTC ACC TTT GGC Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu Val Thr Phe Gly 265 270 275	1230
CAT GAT GGC CGG GGC CAT GCC TTG ACC CGA CGC CGG AGG GCC AAG CGT His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg Ala Lys Arg 280 285 290	1278
AGC CCT AAG CAT CAC TCA CAG CGG GCC AGG AAG AAG AAT AAG AAC TGC Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys 295 300 305	1326
CGG CGC CAC TCG CTC TAT GTG GAC TTC AGC GAT GTG GGC TGG AAT GAC Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp 310 315 320	1374
TGG ATT GTG GCC CCA CCA GGC TAC CAG GCC TTC TAC TGC CAT GGG GAC Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His Gly Asp 325 330 335 340	1422
TGC CCC TTT CCA CTG GCT GAC CAC CTC AAC TCA ACC AAC CAT GCC ATT Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile 345 350 355	1470
GTG CAG ACC CTG GTC AAT TCT GTC AAT TCC AGT ATC CCC AAA GCC TGT Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys Ala Cys 360 365 370	1518
TGT GTG CCC ACT GAA CTG AGT GCC ATC TCC ATG CTG TAC CTG GAT GAG Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu 375 380 385	1566
TAT GAT AAG GTG GTA CTG AAA AAT TAT CAG GAG ATG GTA GTA GAG GGA Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val Glu Gly 390 395 400	1614
TGT GGG TGC CGC TGAGATCAGG CAGTCCTTGA GGATAGACAG ATATACACAC Cys Gly Cys Arg 405	1666
CACACACACA CACCACATAC ACCACACACA CACGTTCCCA TCCACTCACC CACACACTAC ACAGACTGCT TCCTTATAGC TGGACTTTA TTTAAAAAAA AAAAAAAA AATGGAAAAA ATCCCTAAAC ATTACACCTTG ACCTTATTAA TGACTTTACG TGCAAATGTT TTGACCATAT TGATCATATA TTTGACAAA ATATATTAT AACTACGTAT TAAAAGAAAA AAATAAAATG AGTCATTATT TTAAAAAAA AAAAAAAACT CTAGAGTCGA CGGAATTC	1726 1786 1846 1906 1954

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 408 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ile Pro Gly Asn Arg Met Leu Met Val Val Leu Leu Cys Gln Val
1 5 10 15

Leu Leu Gly Gly Ala Ser His Ala Ser Leu Ile Pro Glu Thr Gly Lys
20 25 30

Lys Lys Val Ala Glu Ile Gln Gly His Ala Gly Gly Arg Arg Ser Gly
35 40 45

Gln Ser His Glu Leu Leu Arg Asp Phe Glu Ala Thr Leu Leu Gln Met
50 55 60

Phe Gly Leu Arg Arg Arg Pro Gln Pro Ser Lys Ser Ala Val Ile Pro
65 70 75 80

Asp Tyr Met Arg Asp Leu Tyr Arg Leu Gln Ser Gly Glu Glu Glu
85 90 95

Glu Gln Ile His Ser Thr Gly Leu Glu Tyr Pro Glu Arg Pro Ala Ser
100 105 110

Arg Ala Asn Thr Val Arg Ser Phe His His Glu Glu His Leu Glu Asn
115 120 125

Ile Pro Gly Thr Ser Glu Asn Ser Ala Phe Arg Phe Leu Phe Asn Leu
130 135 140

Ser Ser Ile Pro Glu Asn Glu Val Ile Ser Ser Ala Glu Leu Arg Leu
145 150 155 160

Phe Arg Glu Gln Val Asp Gln Gly Pro Asp Trp Glu Arg Gly Phe His
165 170 175

Arg Ile Asn Ile Tyr Glu Val Met Lys Pro Pro Ala Glu Val Val Pro
180 185 190

Gly His Leu Ile Thr Arg Leu Leu Asp Thr Arg Leu Val His His Asn
195 200 205

Val Thr Arg Trp Glu Thr Phe Asp Val Ser Pro Ala Val Leu Arg Trp
210 215 220

Thr Arg Glu Lys Gln Pro Asn Tyr Gly Leu Ala Ile Glu Val Thr His
225 230 235 240

Leu His Gln Thr Arg Thr His Gln Gly Gln His Val Arg Ile Ser Arg
245 250 255

Ser Leu Pro Gln Gly Ser Gly Asn Trp Ala Gln Leu Arg Pro Leu Leu
260 265 270

Val Thr Phe Gly His Asp Gly Arg Gly His Ala Leu Thr Arg Arg Arg
275 280 285

Arg Ala Lys Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys
290 295 300

Asn	Lys	Asn	Cys	Arg	Arg	His	Ser	Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val
305															320
Gly	Trp	Asn	Asp	Trp	Ile	Val	Ala	Pro	Pro	Gly	Tyr	Gln	Ala	Phe	Tyr
					325					330					335
Cys	His	Gly	Asp	Cys	Pro	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser	Thr
					340			345					350		
Asn	His	Ala	Ile	Val	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Ser	Ile
					355			360				365			
Pro	Lys	Ala	Cys	Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu
					370			375				380			
Tyr	Leu	Asp	Glu	Tyr	Asp	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Glu	Met
					385			390			395				400
Val	Val	Glu	Gly	Cys	Gly	Cys	Cys	Arg							
								405							

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1448 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 97..1389

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTGACCGAGC	GGCGCGGACG	GCCGCCTGCC	CCCTCTGCCA	CCTGGGGCGG	TGCGGGCCCG	60										
GAGCCCGGAG	CCCGGGTAGC	GCGTAGAGCC	GGCGCG	ATG	CAC	GTG	CGC	TCA	CTG	114						
						Met	His	Val	Arg	Ser	Leu					
						1					5					
CGA	GCT	GCG	GCG	CCG	CAC	AGC	TTC	GTG	GCG	CTC	TGG	GCA	CCC	CTG	TTC	162
Arg	Ala	Ala	Ala	Pro	His	Ser	Phe	Val	Ala	Leu	Trp	Ala	Pro	Leu	Phe	
							10		15			20				
CTG	CTG	CGC	TCC	GCC	GAC	TTC	AGC	CTG	GAC	AAC	GAG	GTG	CAC	210		
Leu	Leu	Arg	Ser	Ala	Leu	Ala	Asp	Phe	Ser	Leu	Asp	Asn	Glu	Val	His	
							25		30			35				
TCG	AGC	TTC	ATC	CAC	CGG	CGC	CTC	CGC	AGC	CAG	GAG	CGG	CGG	GAG	ATG	258
Ser	Ser	Phe	Ile	His	Arg	Arg	Leu	Arg	Ser	Gln	Glu	Arg	Arg	Glu	Met	
							40		45			50				
CAG	CGC	GAG	ATC	CTC	TCC	ATT	TTG	GGC	TTG	CCC	CAC	CGC	CCG	CGC	CCG	306
Gln	Arg	Glu	Ile	Leu	Ser	Ile	Leu	Gly	Leu	Pro	His	Arg	Pro	Arg	Pro	
							55		60			65			70	

CAC CTC CAG GGC AAG CAC AAC TCG GCA CCC ATG TTC ATG CTG GAC CTG				354
His Leu Gln Gly Lys His Asn Ser Ala Pro Met Phe Met Leu Asp Leu				
75	80	85		
TAC AAC GCC ATG GCG GTG GAG GAG GGC GGC GGG CCC GGC GGC CAG GGC				402
Tyr Asn Ala Met Ala Val Glu Glu Gly Gly Pro Gly Gly Gln Gly				
90	95	100		
TTC TCC TAC CCC TAC AAG GCC GTC TTC AGT ACC CAG GGC CCC CCT CTG				450
Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser Thr Gln Gly Pro Pro Leu				
105	110	115		
GCC AGC CTG CAA GAT AGC CAT TTC CTC ACC GAC GCC GAC ATG GTC ATG				498
Ala Ser Leu Gln Asp Ser His Phe Leu Thr Asp Ala Asp Met Val Met				
120	125	130		
AGC TTC GTC AAC CTC GTG GAA CAT GAC AAG GAA TTC TTC CAC CCA CGC				546
Ser Phe Val Asn Leu Val Glu His Asp Lys Glu Phe Phe His Pro Arg				
135	140	145	150	
TAC CAC CAT CGA GAG TTC CGG TTT GAT CTT TCC AAG ATC CCA GAA GGG				594
Tyr His His Arg Glu Phe Arg Phe Asp Leu Ser Lys Ile Pro Glu Gly				
155	160	165		
GAA GCT GTC ACG GCA GCC GAA TTC CGG ATC TAC AAG GAC TAC ATC CGG				642
Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Tyr Ile Arg				
170	175	180		
GAA CGC TTC GAC AAT GAG ACG TTC CGG ATC AGC GTT TAT CAG GTG CTC				690
Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile Ser Val Tyr Gln Val Leu				
185	190	195		
CAG GAG CAC TTG GGC AGG GAA TCG GAT CTC TTC CTG CTC GAC AGC CGT				738
Gln Glu His Leu Gly Arg Glu Ser Asp Leu Phe Leu Leu Asp Ser Arg				
200	205	210		
ACC CTC TGG GCC TCG GAG GAG GGC TGG CTG GTG TTT GAC ATC ACA GCC				786
Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu Val Phe Asp Ile Thr Ala				
215	220	225	230	
ACC AGC AAC CAC TGG GTG GTC AAT CCG CGG CAC AAC CTG GGC CTG CAG				834
Thr Ser Asn His Trp Val Val Asn Pro Arg His Asn Leu Gly Leu Gln				
235	240	245		
CTC TCG GTG GAG ACG CTG GAT GGG CAG AGC ATC AAC CCC AAG TTG GCG				882
Leu Ser Val Glu Thr Leu Asp Gly Gln Ser Ile Asn Pro Lys Leu Ala				
250	255	260		
GGC CTG ATT GGG CGG CAC GGG CCC CAG AAC AAG CAG CCC TTC ATG GTG				930
Gly Leu Ile Gly Arg His Gly Pro Gln Asn Lys Gln Pro Phe Met Val				
265	270	275		
GCT TTC TTC AAG GCC ACG GAG GTC CAC TTC CGC AGC ATC CGG TCC ACG				978
Ala Phe Phe Lys Ala Thr Glu Val His Phe Arg Ser Ile Arg Ser Thr				
280	285	290		
GGG AGC AAA CAG CGC AGC CAG AAC CGC TCC AAG ACG CCC AAG AAC CAG				1026
Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser Lys Thr Pro Lys Asn Gln				
295	300	305	310	

CGA GCC CTG CGG ATG GCC AAC GTG GCA GAG AAC AGC AGC AGC GAC CAG Glu Ala Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser Ser Asp Gln 315 320 325	1074
AGG CAG GCC TGT AAG AAG CAC GAG CTG TAT GTC AGC TTC CGA GAC CTG Arg Gln Ala Cys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu 330 335 340	1122
GGC TGG CAG GAC TGG ATC ATC GCG CCT GAA GGC TAC GCC GCC TAC TAC Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr 345 350 355	1170
TGT GAG GGG GAG TGT GCC TTC CCT CTG AAC TCC TAC ATG AAC GCC ACC Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr 360 365 370	1218
AAC CAC GCC ATC GTG CAG ACG CTG GTC CAC TTC ATC AAC CCG GAA ACG Asn His Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr 375 380 385 390	1266
GTG CCC AAG CCC TGC TGT GCG CCC ACG CAG CTC AAT GCC ATC TCC GTC Val Pro Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val 395 400 405	1314
CTC TAC TTC GAT GAC AGC TCC AAC GTC ATC CTG AAG AAA TAC AGA AAC Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn 410 415 420	1362
ATG GTG GTC CGG GCC TGT GGC TGC CAC TAGCTCCTCC GAGAATTCA Met Val Val Arg Ala Cys Gly Cys His 425 430	1409
ACCCCTTTGGGG GCCAAGTTTT TCTGGATCCT CCATTGCTC	1448

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 431 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met His Val Arg Ser Leu Arg Ala Ala Ala Pro His Ser Phe Val Ala
1 5 10 15
Leu Trp Ala Pro Leu Phe Leu Leu Arg Ser Ala Leu Ala Asp Phe Ser
20 25 30
Leu Asp Asn Glu Val His Ser Ser Phe Ile His Arg Arg Leu Arg Ser
35 40 45
Gln Glu Arg Arg Glu Met Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu
50 55 60
Pro His Arg Pro Arg Pro His Leu Gln Gly Lys His Asn Ser Ala Pro

```

65 70 75 80

Met Phe Met Leu Asp Leu Tyr Asn Ala Met Ala Val Glu Glu Gly Gly
85 90 95

Gly Pro Gly Gly Gln Gly Phe Ser Tyr Pro Tyr Lys Ala Val Phe Ser
100 105 110

Thr Gln Gly Pro Pro Leu Ala Ser Leu Gln Asp Ser His Phe Leu Thr
115 120 125

Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu His Asp Lys
130 135 140

Glu Phe Phe His Pro Arg Tyr His His Arg Glu Phe Arg Phe Asp Leu
145 150 155 160

Ser Lys Ile Pro Glu Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile
165 170 175

Tyr Lys Asp Tyr Ile Arg Glu Arg Phe Asp Asn Glu Thr Phe Arg Ile
180 185 190

Ser Val Tyr Gln Val Leu Gln Glu His Leu Gly Arg Glu Ser Asp Leu
195 200 205

Phe Leu Leu Asp Ser Arg Thr Leu Trp Ala Ser Glu Glu Gly Trp Leu
210 215 220

Val Phe Asp Ile Thr Ala Thr Ser Asn His Trp Val Val Asn Pro Arg
225 230 235 240

His Asn Leu Gly Leu Gln Leu Ser Val Glu Thr Leu Asp Gly Gln Ser
245 250 255

Ile Asn Pro Lys Leu Ala Gly Leu Ile Gly Arg His Gly Pro Gln Asn
260 265 270

Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Thr Glu Val His Phe
275 280 285

Arg Ser Ile Arg Ser Thr Gly Ser Lys Gln Arg Ser Gln Asn Arg Ser
290 295 300

Lys Thr Pro Lys Asn Gln Glu Ala Leu Arg Met Ala Asn Val Ala Glu
305 310 315 320

Asn Ser Ser Ser Asp Gln Arg Gln Ala Cys Lys Lys His Glu Leu Tyr
325 330 335

Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Glu
340 345 350

Gly Tyr Ala Ala Tyr Tyr Cys Glu Gly Glu Cys Ala Phe Pro Leu Asn
355 360 365

Ser Tyr Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His
370 375 380

Phe Ile Asn Pro Glu Thr Val Pro Lys Pro Cys Cys Ala Pro Thr Gln

385

390

395

400

Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn Val Ile
 405 410 415

Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
 420 425 430

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2923 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (F) TISSUE TYPE: Human placenta

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Stratagene catalog #936203 Human placenta
 cDNA library
- (B) CLONE: BMP6C35

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 160..1701

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 1282..1698

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 1..2923

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACCATGAG AGATAAGGAC TGAGGGCCAG GAAGGGGAAG CGAGCCCGCC GAGAGGTGGC	60
GGGGACTGCT CACGCCAAGG GCCACAGCGG CCGCGCTCCG GCCTCGCTCC GCCGCTCCAC	120
GCCTCGCGGG ATCCGGGGGG GCAGCCCGGC CGGGCGGGGG ATG CCG GGG CTG GGG Met Pro Gly Leu Gly -374 -370	174
CGG AGG GCG CAG TGG CTG TGC TGG TGG GGG CTG CTG TGC AGC TGC Arg Arg Ala Gln Trp Leu Cys Trp Trp Gly Leu Leu Cys Ser Cys -365 -360 -355	222
TGC GGG CCC CCG CCG CTG CGG CCG CCC TTG CCC GCT GCC GCG GCC GCC	270

Cys	Gly	Pro	Pro	Leu	Arg	Pro	Pro	Leu	Pro	Ala	Ala	Ala	Ala	Ala		
		-350				-345				-340						
GCC	GCC	GGG	GGG	CAG	CTG	CTG	GGG	GAC	GGC	GGG	AGC	CCC	GGC	CGC	ACG	318
Ala	Ala	Gly	Gly	Gln	Leu	Leu	Gly	Asp	Gly	Gly	Ser	Pro	Gly	Arg	Thr	
		-335				-330				-325						
GAG	CAG	CCG	CCG	CCG	TCG	CCG	CAG	TCC	TCC	TCG	GGC	TTC	CTG	TAC	CGG	366
Glu	Gln	Pro	Pro	Pro	Ser	Pro	Gln	Ser	Ser	Gly	Phe	Leu	Tyr	Arg		
		-320				-315				-310						
CGG	CTC	AAG	ACG	CAG	GAG	AAG	CGG	GAG	ATG	CAG	AAG	GAG	ATC	TTG	TCG	414
Arg	Leu	Lys	Thr	Gln	Glu	Lys	Arg	Glu	Met	Gln	Lys	Glu	Ile	Leu	Ser	
		-305				-300				-295			-290			
GTG	CTG	GGG	CTC	CCG	CAC	CGG	CCC	CGG	CCC	CTG	CAC	GGC	CTC	CAA	CAG	462
Val	Leu	Gly	Leu	Pro	His	Arg	Pro	Arg	Pro	Leu	His	Gly	Leu	Gln	Gln	
		-285				-280				-275						
CCG	CAG	CCC	CCG	GCG	CTC	CGG	CAG	CAG	GAG	CAG	CAG	CAG	CAG	CAG	CAG	510
Pro	Gln	Pro	Pro	Ala	Leu	Arg	Gln	Gln	Glu	Glu	Gln	Gln	Gln	Gln	Gln	
		-270				-265				-260						
CAG	CTG	CCT	CGC	GGA	GAG	CCC	CCT	CCC	GGG	CGA	CTG	AAG	TCC	GCG	CCC	558
Gln	Leu	Pro	Arg	Gly	Glu	Pro	Pro	Pro	Gly	Arg	Leu	Lys	Ser	Ala	Pro	
		-255				-250				-245						
CTC	TTC	ATG	CTG	GAT	CTG	TAC	AAC	GCC	CTG	TCC	GCC	GAC	AAC	GAC	GAG	606
Leu	Phe	Met	Leu	Asp	Leu	Tyr	Asn	Ala	Leu	Ser	Ala	Asp	Asn	Asp	Glu	
		-240				-235				-230						
GAC	GGG	GCG	TCG	GAG	GGG	GAG	AGG	CAG	CAG	TCC	TGG	CCC	CAC	GAA	GCA	654
Asp	Gly	Ala	Ser	Glu	Gly	Glu	Arg	Gln	Gln	Ser	Trp	Pro	His	Glu	Ala	
		-225				-220				-215			-210			
GCC	AGC	TCG	TCC	CAG	CGT	CGG	CAG	CCC	CCG	GGC	GCC	GCG	CAC	CCG		702
Ala	Ser	Ser	Ser	Gln	Arg	Arg	Gln	Pro	Pro	Pro	Gly	Ala	Ala	His	Pro	
		-205				-200				-195						
CTC	AAC	CGC	AAG	AGC	CTT	CTG	GCC	CCC	GGA	TCT	GGC	AGC	GGC	GGC	GCG	750
Leu	Asn	Arg	Lys	Ser	Leu	Leu	Ala	Pro	Gly	Ser	Gly	Ser	Gly	Gly	Ala	
		-190				-185				-180						
TCC	CCA	CTG	ACC	AGC	GCG	CAG	GAC	AGC	GCC	TTC	CTC	AAC	GAC	GCG	GAC	798
Ser	Pro	Leu	Thr	Ser	Ala	Gln	Asp	Ser	Ala	Phe	Leu	Asn	Asp	Ala	Asp	
		-175				-170				-165						
ATG	GTC	ATG	AGC	TTT	GTG	AAC	CTG	GTG	GAG	TAC	GAC	AAG	GAG	TTC	TCC	846
Met	Val	Met	Ser	Phe	Val	Asn	Leu	Val	Glu	Tyr	Asp	Lys	Glu	Phe	Ser	
		-160				-155				-150						
CCT	CGT	CAG	CGA	CAC	CAC	AAA	GAG	TTC	AAG	TTC	AAC	TTA	TCC	CAG	ATT	894
Pro	Arg	Gln	Arg	His	His	Lys	Glu	Phe	Lys	Phe	Asn	Leu	Ser	Gln	Ile	
		-145				-140				-135			-130			
CCT	GAG	GGT	GAG	GTG	GTG	ACG	GCT	GCA	GAA	TTC	CGC	ATC	TAC	AAG	GAC	942
Pro	Glu	Gly	Glu	Val	Val	Thr	Ala	Ala	Glu	Phe	Arg	Ile	Tyr	Lys	Asp	
		-125				-120				-115						
TGT	GTT	ATG	GGG	AGT	TTT	AAA	AAC	CAA	ACT	TTT	CTT	ATC	AGC	ATT	TAT	990

Cys Val Met Gly Ser Phe Lys Asn Gln Thr Phe Leu Ile Ser Ile Tyr			
-110	-105	-100	
CAA GTC TTA CAG GAG CAT CAG CAC AGA GAC TCT GAC CTG TTT TTG TTG			1038
Gln Val Leu Gln Glu His Gln His Arg Asp Ser Asp Leu Phe Leu Leu			
-95	-90	-85	
GAC ACC CGT GTA GTA TGG GCC TCA GAA GAA GGC TGG CTG GAA TTT GAC			1086
Asp Thr Arg Val Val Trp Ala Ser Glu Glu Gly Trp Leu Glu Phe Asp			
-80	-75	-70	
ATC ACG GCC ACT AGC AAT CTG TGG GTT GTG ACT CCA CAG CAT AAC ATG			1134
Ile Thr Ala Thr Ser Asn Leu Trp Val Val Thr Pro Gln His Asn Met			
-65	-60	-55	-50
GGG CTT CAG CTG AGC GTG GTG ACA AGG GAT GGA GTC CAC GTC CAC CCC			1182
Gly Leu Gln Leu Ser Val Val Thr Arg Asp Gly Val His Val His Pro			
-45	-40	-35	
CGA GCC GCA GGC CTG GTG GGC AGA GAC GGC CCT TAC GAT AAG CAG CCC			1230
Arg Ala Ala Gly Leu Val Gly Arg Asp Gly Pro Tyr Asp Lys Gln Pro			
-30	-25	-20	
TTC ATG GTG GCT TTC TTC AAA GTG AGT GAG GTC CAC GTG CGC ACC ACC			1278
Phe Met Val Ala Phe Phe Lys Val Ser Glu Val His Val Arg Thr Thr			
-15	-10	-5	
AGG TCA GCC TCC AGC CGG CGC CGA CAA CAG AGT CGT AAT CGC TCT ACC			1326
Arg Ser Ala Ser Ser Arg Arg Gln Gln Ser Arg Asn Arg Ser Thr			
1	5	10	15
CAG TCC CAG GAC GTG GCG CGG GTC TCC AGT GCT TCA GAT TAC AAC AGC			1374
Gln Ser Gln Asp Val Ala Arg Val Ser Ser Ala Ser Asp Tyr Asn Ser			
20	25	30	
AGT GAA TTG AAA ACA GCC TGC AGG AAG CAT GAG CTG TAT GTG AGT TTC			1422
Ser Glu Leu Lys Thr Ala Cys Arg Lys His Glu Leu Tyr Val Ser Phe			
35	40	45	
CAA GAC CTG GGA TGG CAG GAC TGG ATC ATT GCA CCC AAG GGC TAT GCT			1470
Gln Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala			
50	55	60	
GCC AAT TAC TGT GAT GGA GAA TGC TCC TTC CCA CTC AAC GCA CAC ATG			1518
Ala Asn Tyr Cys Asp Gly Glu Cys Ser Phe Pro Leu Asn Ala His Met			
65	70	75	
AAT GCA ACC AAC CAC GCG ATT GTG CAG ACC TTG GTT CAC CTT ATG AAC			1566
Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu Val His Leu Met Asn			
80	85	90	95
CCC GAG TAT GTC CCC AAA CCG TGC TGT GCG CCA ACT AAG CTA AAT GCC			1614
Pro Glu Tyr Val Pro Lys Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala			
100	105	110	
ATC TCG GTT CTT TAC TTT GAT GAC AAC TCC AAT GTC ATT CTG AAA AAA			1662
Ile Ser Val Leu Tyr Phe Asp Asp Asn Ser Asn Val Ile Leu Lys Lys			
115	120	125	
TAC AGG AAT ATG GTT GTA AGA GCT TGT GGA TGC CAC TAACTCGAAA			1708

Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys His
 130 135 140

CCAGATGCTG	GGGACACACA	TTCTGCCTTG	GATTCCCTAGA	TTACATCTGC	CTTAAAAAAA	1768
CACGGAAGCA	CAGTTGGAGG	TGGGACGGATG	AGACTTTGAA	ACTATCTCAT	GCCAGTGCCT	1828
TATTACCCAG	GAAGATTTA	AAGGACCTCA	TTAATAATT	GCTCACTTGG	TAAATGACGT	1888
GAGTAGTTGT	TGGTCTGTAG	CAAGCTGAGT	TTGGATGTCT	GTAGCATAAG	GTCTGGTAAC	1948
TGCAGAAACA	TAACCGTGAA	GCTCTCCTA	CCCTCCTCCC	CCAAAAACCC	ACCAAAATTA	2008
GTGTTAGCTG	TAGATCAAGC	TATTTGGGT	GTGTTAGT	AAATAGGGAA	AATAATCTCA	2068
AAGGAGTTAA	ATGTATTCTT	GGCTAAAGGA	TCAGCTGGTT	CAGTACTGTC	TATCAAAGGT	2128
AGATTTTACA	GAGAACAGAA	ATCGGGGAAG	TGGGGGGAAC	GCCTCTGTT	AGTTCAATTCC	2188
CAGAAGTCCA	CAGGACGCAC	AGCCCAGGCC	ACAGCCAGGG	CTCCACGGGG	CGCCCTTGTC	2248
TCAGTCATTG	CTGTTGTATG	TTCGTGCTGG	AGTTTGTTG	GTGTAAAAT	ACACTTATTT	2308
CAGCCAAAAC	ATACCATTTC	TACACCTCAA	TCCTCCATT	GCTGTACTCT	TTGCTAGTAC	2368
CAAAAGTAGA	CTGATTACAC	TGAGGTGAGG	CTACAAGGGG	TGTGTAACCG	TGTAACACGT	2428
GAAGGCAGTG	CTCACCTCTT	CTTTACCAGA	ACGGTTCTTT	GACCAGCACA	TTAACTTCTG	2488
GAAGTCCGGC	TCTAGTACCT	TTTCAGTAAA	GTGGTTCTCT	GCCTTTTAC	TATACAGCAT	2548
ACCACGCCAC	AGGGTTAGAA	CCAACGAAGA	AAATAAAATG	AGGGTGCCCA	GCTTATAAGA	2608
ATGGTGTAG	GGGGATGAGC	ATGCTGTTA	TGAACGGAAA	TCATGATTTC	CCTGTAGAAA	2668
GTGAGGCTCA	GATTAAATTT	TAGAATATT	TCTAAATGTC	TTTTCACAA	TCATGTGACT	2728
GGGAAGGCAA	TTTCATACTA	AACTGATTAA	ATAATACATT	TATAATCTAC	AACTGTTGC	2788
ACTTACAGCT	TTTTTGTA	ATATAAACTA	TAATTTATTG	TCTATTAT	ATCTGTTTG	2848
CTGTGGCGTT	GGGGGGGGGG	CCGGGCTTTT	GGGGGGGGGG	GTGGTTGG	GGGGTGTGCGT	2908
GGTGTGGCGG	GGCGG					2923

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 513 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Pro Gly Leu Gly Arg Arg Ala Gln Trp Leu Cys Trp Trp Trp Gly
 -374 -370 -365 -360

Leu Leu Cys Ser Cys Cys Gly Pro Pro Pro Leu Arg Pro Pro Leu Pro
-355 -350 -345

Ala Ala Ala Ala Ala Ala Gly Gly Gln Leu Leu Gly Asp Gly Gly
-340 -335 -330

Ser Pro Gly Arg Thr Glu Gln Pro Pro Pro Ser Pro Gln Ser Ser Ser
-325 -320 -315

Gly Phe Leu Tyr Arg Arg Leu Lys Thr Gln Glu Lys Arg Glu Met Gln
-310 -305 -300 -295

Lys Glu Ile Leu Ser Val Leu Gly Leu Pro His Arg Pro Arg Pro Leu
-290 -285 -280

His Gly Leu Gln Gln Pro Gln Pro Pro Ala Leu Arg Gln Gln Glu Glu
-275 -270 -265

Gln Gln Gln Gln Gln Leu Pro Arg Gly Glu Pro Pro Pro Gly Arg
-260 -255 -250

Leu Lys Ser Ala Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Leu Ser
-245 -240 -235

Ala Asp Asn Asp Glu Asp Gly Ala Ser Glu Gly Glu Arg Gln Gln Ser
-230 -225 -220 -215

Trp Pro His Glu Ala Ala Ser Ser Ser Gln Arg Arg Gln Pro Pro Pro
-210 -205 -200

Gly Ala Ala His Pro Leu Asn Arg Lys Ser Leu Leu Ala Pro Gly Ser
-195 -190 -185

Gly Ser Gly Gly Ala Ser Pro Leu Thr Ser Ala Gln Asp Ser Ala Phe
-180 -175 -170

Leu Asn Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu Tyr
-165 -160 -155

Asp Lys Glu Phe Ser Pro Arg Gln Arg His His Lys Glu Phe Lys Phe
-150 -145 -140 -135

Asn Leu Ser Gln Ile Pro Glu Gly Glu Val Val Thr Ala Ala Glu Phe
-130 -125 -120

Arg Ile Tyr Lys Asp Cys Val Met Gly Ser Phe Lys Asn Gln Thr Phe
-115 -110 -105

Leu Ile Ser Ile Tyr Gln Val Leu Gln Glu His Gln His Arg Asp Ser
-100 -95 -90

Asp Leu Phe Leu Leu Asp Thr Arg Val Val Trp Ala Ser Glu Glu Gly
-85 -80 -75

Trp Leu Glu Phe Asp Ile Thr Ala Thr Ser Asn Leu Trp Val Val Thr
-70 -65 -60 -55

Pro Gln His Asn Met Gly Leu Gln Leu Ser Val Val Thr Arg Asp Gly
-50 -45 -40

Val His Val His Pro Arg Ala Ala Gly Leu Val Gly Arg Asp Gly Pro
-35 -30 -25

Tyr Asp Lys Gln Pro Phe Met Val Ala Phe Phe Lys Val Ser Glu Val
-20 -15 -10

His Val Arg Thr Thr Arg Ser Ala Ser Ser Arg Arg Arg Gln Gln Ser
-5 1 5 10

Arg Asn Arg Ser Thr Gln Ser Gln Asp Val Ala Arg Val Ser Ser Ala
15 20 25

Ser Asp Tyr Asn Ser Ser Glu Leu Lys Thr Ala Cys Arg Lys His Glu
30 35 40

Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala
45 50 55

Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp Gly Glu Cys Ser Phe Pro
60 65 70

Leu Asn Ala His Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu
75 80 85 90

Val His Leu Met Asn Pro Glu Tyr Val Pro Lys Pro Cys Cys Ala Pro
95 100 105

Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Asn Ser Asn
110 115 120

Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ala Cys Gly Cys
125 130 135

His

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2153 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (iii) HYPOTHETICAL: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Homo sapiens
 - (H) CELL LINE: U2-OS osteosarcoma
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: U2-OS human osteosarcoma cDNA library
 - (B) CLONE: U2-16
- (viii) POSITION IN GENOME:
 - (C) UNITS: bp
- (ix) FEATURE:
 - (A) NAME/KEY: CDS

(B) LOCATION: 699..2063

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 1647..2060

(ix) FEATURE:

(A) NAME/KEY: mRNA
(B) LOCATION: 1..2153

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGTATATT TGTGCCTGCT GGAGGTGGAA TTAACAGTAA GAAGGGAGAAA GGGATTGAAT	60
GGACTTACAG GAAGGGATTTC AAGTAAATTC AGGGAAACAC ATTTACTTGA ATAGTACAAC	120
CTAGAGTATT ATTTTACACT AAGACGACAC AAAAGATGTT AAAGTTATCA CCAAGCTGCC	180
GGACAGATAT ATATTCCAAC ACCAAGGTGC AGATCAGCAT AGATCTGTGA TTCAGAAATC	240
AGGATTGTT TTGGAAAGAG CTCAGGGTT GAGAAGAACT CAAAAGCAAG TGAAGATTAC	300
TTTGGGAACT ACAGTTTATC AGAAGATCAA CTTTGCTAA TTCAAATACC AAAGGCCTGA	360
TTATCATAAA TTCATATAGG AATGCATAGG TCATCTGATC AAATAATATT AGCCGTCTTC	420
TGCTACATCA ATGCAGCAAA AACTCTAAC AACTGTGGAT AATTGGAAAT CTGAGTTCA	480
GCTTCTTAG AAATAACTAC TCTTGACATA TTCCAAAATA TTTAAAATAG GACAGGAAAA	540
TCGGTGAGGA TGTTGTGCTC AGAAATGTCA CTGTCATGAA AAATAGGTAA ATTTGTTTT	600
TCAGCTACTG GGAAACTGTA CCTCCTAGAA CCTTAGGTTT TTTTTTTTT AAGAGGACAA	660
GAAGGACTAA AAATATCAAC TTTTGCTTTT GGACAAAAA ATG CAT CTG ACT GTA	713
Met His Leu Thr Val	
-316-315	
TTT TTA CTT AAG GGT ATT GTG GGT TTC CTC TGG AGC TGC TGG GTT CTA	761
Phe Leu Leu Lys Gly Ile Val Gly Phe Leu Trp Ser Cys Trp Val Leu	
-310 -305 -300	
GTG GGT TAT GCA AAA GGA GGT TTG GGA GAC AAT CAT GTT CAC TCC AGT	809
Val Gly Tyr Ala Lys Gly Leu Gly Asp Asn His Val His Ser Ser	
-295 -290 -285 -280	
TTT ATT TAT AGA AGA CTA CGG AAC CAC GAA AGA CGG GAA ATA CAA AGG	857
Phe Ile Tyr Arg Arg Leu Arg Asn His Glu Arg Arg Glu Ile Gln Arg	
-275 -270 -265	
GAA ATT CTC TCT ATC TTG GGT TTG CCT CAC AGA CCC AGA CCA TTT TCA	905
Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg Pro Arg Pro Phe Ser	
-260 -255 -250	
CCT GGA AAA ATG ACC AAT CAA GCG TCC TCT GCA CCT CTC TTT ATG CTG	953
Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala Pro Leu Phe Met Leu	
-245 -240 -235	
GAT CTC TAC AAT GCC GAA GAA AAT CCT GAA GAG TCG GAG TAC TCA GTA	1001

Asp Leu Tyr Asn Ala Glu Glu Asn Pro Glu Glu Ser Glu Tyr Ser Val		
-230	-225	-220
AGG GCA TCC TTG GCA GAA GAG ACC AGA GGG GCA AGA AAG GGA TAC CCA		1049
Arg Ala Ser Leu Ala Glu Glu Thr Arg Gly Ala Arg Lys Gly Tyr Pro		
-215	-210	-205
-200		
GCC TCT CCC AAT GGG TAT CCT CGT CGC ATA CAG TTA TCT CGG ACG ACT		1097
Ala Ser Pro Asn Gly Tyr Pro Arg Arg Ile Gln Leu Ser Arg Thr Thr		
-195	-190	-185
CCT CTG ACC ACC CAG AGT CCT CCT CTA GCC AGC CTC CAT GAT ACC AAC		1145
Pro Leu Thr Thr Gln Ser Pro Pro Leu Ala Ser Leu His Asp Thr Asn		
-180	-175	-170
TTT CTG AAT GAT GCT GAC ATG GTC ATG AGC TTT GTC AAC TTA GTT GAA		1193
Phe Leu Asn Asp Ala Asp Met Val Met Ser Phe Val Asn Leu Val Glu		
-165	-160	-155
AGA GAC AAG GAT TTT TCT CAC CAG CGA AGG CAT TAC AAA GAA TTT CGA		1241
Arg Asp Lys Asp Phe Ser His Gln Arg Arg His Tyr Lys Glu Phe Arg		
-150	-145	-140
TTT GAT CTT ACC CAA ATT CCT CAT GGA GAG GCA GTG ACA GCA GCT GAA		1289
Phe Asp Leu Thr Gln Ile Pro His Gly Glu Ala Val Thr Ala Ala Glu		
-135	-130	-125
-120		
TTC CGG ATA TAC AAG GAC CGG AGC AAC CGA TTT GAA AAT GAA ACA		1337
Phe Arg Ile Tyr Lys Asp Arg Ser Asn Asn Arg Phe Glu Asn Glu Thr		
-115	-110	-105
ATT AAG ATT AGC ATA TAT CAA ATC ATC AAG GAA TAC ACA AAT AGG GAT		1385
Ile Lys Ile Ser Ile Tyr Gln Ile Ile Lys Glu Tyr Thr Asn Arg Asp		
-100	-95	-90
GCA GAT CTG TTC TTG TTA GAC ACA AGA AAG GCC CAA GCT TTA GAT GTG		1433
Ala Asp Leu Phe Leu Leu Asp Thr Arg Lys Ala Gln Ala Leu Asp Val		
-85	-80	-75
GGT TGG CTT GTC TTT GAT ATC ACT GTG ACC AGC AAT CAT TGG GTG ATT		1481
Gly Trp Leu Val Phe Asp Ile Thr Val Thr Ser Asn His Trp Val Ile		
-70	-65	-60
AAT CCC CAG AAT AAT TTG GGC TTA CAG CTC TGT GCA GAA ACA GGG GAT		1529
Asn Pro Gln Asn Asn Leu Gly Leu Gln Leu Cys Ala Glu Thr Gly Asp		
-55	-50	-45
-40		
GGA CGC AGT ATC AAC GTA AAA TCT GCT GGT CTT GTG GGA AGA CAG GGA		1577
Gly Arg Ser Ile Asn Val Lys Ser Ala Gly Leu Val Gly Arg Gln Gly		
-35	-30	-25
CCT CAG TCA AAA CAA CCA TTC ATG GTG GCC TTC AAG GCG AGT GAG		1625
Pro Gln Ser Lys Gln Pro Phe Met Val Ala Phe Phe Lys Ala Ser Glu		
-20	-15	-10
GTA CTT CTT CGA TCC GTG AGA GCA GCC AAC AAA CGA AAA AAT CAA AAC		1673
Val Leu Leu Arg Ser Val Arg Ala Ala Asn Lys Arg Lys Asn Gln Asn		
-5	1	5
CGC AAT AAA TCC AGC TCT CAT CAG GAC TCC TCC AGA ATG TCC AGT GTT		1721

Arg Asn Lys Ser Ser Ser His Gln Asp Ser Ser Arg Met Ser Ser Val		
10 15 20 25		
GGA GAT TAT AAC ACA AGT GAG CAA AAA CAA GCC TGT AAG AAG CAC GAA		1769
Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln Ala Cys Lys Lys His Glu		
30 35 40		
CTC TAT GTG AGC TTC CGG GAT CTG GGA TGG CAG GAC TGG ATT ATA GCA		1817
Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln Asp Trp Ile Ile Ala		
45 50 55		
CCA GAA GGA TAC GCT GCA TTT TAT TGT GAT GGA GAA TGT TCT TTT CCA		1865
Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly Glu Cys Ser Phe Pro		
60 65 70		
CTT AAC GCC CAT ATG AAT GCC ACC AAC CAC GCT ATA GTT CAG ACT CTG		1913
Leu Asn Ala His Met Asn Ala Thr Asn His Ala Ile Val Gln Thr Leu		
75 80 85		
GTT CAT CTG ATG TTT CCT GAC CAC GTA CCA AAG CCT TGT TGT GCT CCA		1961
Val His Leu Met Phe Pro Asp His Val Pro Lys Pro Cys Cys Ala Pro		
90 95 100 105		
ACC AAA TTA AAT GCC ATC TCT GTT CTG TAC TTT GAT GAC AGC TCC AAT		2009
Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe Asp Asp Ser Ser Asn		
110 115 120		
GTC ATT TTG AAA AAA TAT AGA AAT ATG GTA GTA CGC TCA TGT GGC TGC		2057
Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val Arg Ser Cys Gly Cys		
125 130 135		
CAC TAATATTAAA TAATATTGAT AATAACAAAA AGATCTGTAT TAAGGTTTAT		2110
His		
GGCTGCAATA AAAAGCATAAC TTTCAGACAA ACAGAAAAAA AAA		2153

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 454 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met His Leu Thr Val Phe Leu Leu Lys Gly Ile Val Gly Phe Leu Trp
-316 -315 -310 -305

Ser Cys Trp Val Leu Val Gly Tyr Ala Lys Gly Gly Leu Gly Asp Asn
-300 -295 -290 -285

His Val His Ser Ser Phe Ile Tyr Arg Arg Leu Arg Asn His Glu Arg
-280 -275 -270

Arg Glu Ile Gln Arg Glu Ile Leu Ser Ile Leu Gly Leu Pro His Arg
-265 -260 -255

Pro Arg Pro Phe Ser Pro Gly Lys Met Thr Asn Gln Ala Ser Ser Ala
-250 -245 -240

Pro Leu Phe Met Leu Asp Leu Tyr Asn Ala Glu Glu Asn Pro Glu Glu
-235 -230 -225

Ser Glu Tyr Ser Val Arg Ala Ser Leu Ala Glu Glu Thr Arg Gly Ala
-220 -215 -210 -205

Arg Lys Gly Tyr Pro Ala Ser Pro Asn Gly Tyr Pro Arg Arg Ile Gln
-200 -195 -190

Leu Ser Arg Thr Thr Pro Leu Thr Thr Gln Ser Pro Pro Leu Ala Ser
-185 -180 -175

Leu His Asp Thr Asn Phe Leu Asn Asp Ala Asp Met Val Met Ser Phe
-170 -165 -160

Val Asn Leu Val Glu Arg Asp Lys Asp Phe Ser His Gln Arg Arg His
-155 -150 -145

Tyr Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro His Gln Arg Arg His
-140 -135 -130 -125

Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Asp Arg Ser Asn Asn Arg
-120 -115 -110

Phe Glu Asn Glu Thr Ile Lys Ile Ser Ile Tyr Gln Ile Ile Lys Glu
-105 -100 -95

Tyr Thr Asn Arg Asp Ala Asp Leu Phe Leu Leu Asp Thr Arg Lys Ala
-90 -85 -80

Gln Ala Leu Asp Val Gly Trp Leu Val Phe Asp Ile Thr Val Thr Ser
-75 -70 -65

Asn His Trp Val Ile Asn Pro Gln Asn Asn Leu Gly Leu Gln Leu Cys
-60 -55 -50 -45

Ala Glu Thr Gly Asp Gly Arg Ser Ile Asn Val Lys Ser Ala Gly Leu
-40 -35 -30

Val Gly Arg Gln Gly Pro Gln Ser Lys Gln Pro Phe Met Val Ala Phe
-25 -20 -15

Phe Lys Ala Ser Glu Val Leu Leu Arg Ser Val Arg Ala Ala Asn Lys
-10 -5 1

Arg Lys Asn Gln Asn Arg Asn Lys Ser Ser Ser His Gln Asp Ser Ser
5 10 15 20

Arg Met Ser Ser Val Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln Ala
25 30 35

Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp Gln
40 45 50

Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp Gly
55 60 65

Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His Ala
70 75 80

Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro Lys
85 90 95 100

Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr Phe
105 110 115

Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val Val
120 125 130

Arg Ser Cys Gly Cys His
135

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1003 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Homo sapiens
- (F) TISSUE TYPE: Human Heart

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Human heart cDNA library stratagene catalog
#936208
- (B) CLONE: hH38

(viii) POSITION IN GENOME:

- (C) UNITS: bp

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 8..850

(ix) FEATURE:

- (A) NAME/KEY: mat_peptide
- (B) LOCATION: 427..843

(ix) FEATURE:

- (A) NAME/KEY: mRNA
- (B) LOCATION: 1..997

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAATTCC GAG CCC CAT TGG AAG GAG TTC CGC TTT GAC CTG ACC CAG ATC
Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile
-139 -135 -130

CCG GCT GGG GAG GCG GTC ACA GCT GCG GAG TTC CCG ATT TAC AAG GTG Pro Ala Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val -125 -120 -115 -110	97
CCC AGC ATC CAC CTG CTC AAC AGG ACC CTC CAC GTC AGC ATG TTC CAG Pro Ser Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln -105 -100 -95	145
GTG GTC CAG GAG CAG TCC AAC AGG GAG TCT GAC TTG TTC TTT TTG GAT Val Val Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp -90 -85 -80	193
CTT CAG ACG CTC CGA GCT GGA GAC GAG GGC TGG CTG GTG CTG GAT GTC Leu Gln Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val -75 -70 -65	241
ACA GCA GCC AGT GAC TGC TGG TTG CTG AAG CGT CAC AAG GAC CTG GGA Thr Ala Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly -60 -55 -50	289
CTC CGC CTC TAT GTG GAG ACT GAG GAT GGG CAC AGC GTG GAT CCT GGC Leu Arg Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly -45 -40 -35 -30	337
CTG GCC GGC CTG CTG GGT CAA CGG GCC CCA CGC TCC CAA CAG CCT TTC Leu Ala Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe -25 -20 -15	385
GTG GTC ACT TTC TTC AGG GCC AGT CCG AGT CCC ATC CGC ACC CCT CGG Val Val Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg -10 -5 1	433
GCA GTG AGG CCA CTG AGG AGG AGG CAG CCG AAG AAA AGC AAC GAG CTG Ala Val Arg Pro Leu Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu 5 10 15	481
CCG CAG GCC AAC CGA CTC CCA GGG ATC TTT GAT GAC GTC CAC GGC TCC Pro Gln Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser 20 25 30 35	529
CAC GGC CGG CAG GTC TGC CGT CGG CAC GAG CTC TAC GTC AGC TTC CAG His Gly Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln 40 45 50	577
GAC CTT GGC TGG CTG GAC TGG GTC ATC GCC CCC CAA GGC TAC TCA GCC Asp Leu Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala 55 60 65	625
TAT TAC TGT GAG GGG GAG TGC TCC TTC CCG CTG GAC TCC TGC ATG AAC Tyr Tyr Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn 70 75 80	673
GCC ACC AAC CAC GCC ATC CTG CAG TCC CTG GTG CAC CTG ATG AAG CCA Ala Thr Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro 85 90 95	721
AAC GCA GTC CCC AAG GCG TGC TGT GCA CCC ACC AAG CTG AGC GCC ACC Asn Ala Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr 100 105 110 115	769

TCT GTG CTC TAC TAT GAC AGC AGC AAC AAC GTC ATC CTG CGC AAG CAC	817
Ser Val Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His	
120 125 130	
CGC AAC ATG GTG GTC AAG GCC TGC GGC TGC CAC TGAGTCAGCC CGCCCAGCCC	870
Arg Asn Met Val Val Lys Ala Cys Gly Cys His	
135 140	
TACTGCAGCC ACCCTTCTCA TCTGGATCGG GCCCTGCAGA GGCAGAAAAC CCTTAAATGC	930
TGTCACAGCT CAAGCAGGAG TGTCAGGGC CCTCACTCTC GGTGCCTACT TCCTGTCAGG	990
CTTCTGGGAA TTC	1003

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 281 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Glu Pro His Trp Lys Glu Phe Arg Phe Asp Leu Thr Gln Ile Pro Ala
-139 -135 -130 -125

Gly Glu Ala Val Thr Ala Ala Glu Phe Arg Ile Tyr Lys Val Pro Ser
-120 -115 -110

Ile His Leu Leu Asn Arg Thr Leu His Val Ser Met Phe Gln Val Val
-105 -100 -95

Gln Glu Gln Ser Asn Arg Glu Ser Asp Leu Phe Phe Leu Asp Leu Gln
-90 -85 -80

Thr Leu Arg Ala Gly Asp Glu Gly Trp Leu Val Leu Asp Val Thr Ala
-75 -70 -65 -60

Ala Ser Asp Cys Trp Leu Leu Lys Arg His Lys Asp Leu Gly Leu Arg
-55 -50 -45

Leu Tyr Val Glu Thr Glu Asp Gly His Ser Val Asp Pro Gly Leu Ala
-40 -35 -30

Gly Leu Leu Gly Gln Arg Ala Pro Arg Ser Gln Gln Pro Phe Val Val
-25 -20 -15

Thr Phe Phe Arg Ala Ser Pro Ser Pro Ile Arg Thr Pro Arg Ala Val
-10 -5 1 5

Arg Pro Leu Arg Arg Arg Gln Pro Lys Lys Ser Asn Glu Leu Pro Gln
10 15 20

Ala Asn Arg Leu Pro Gly Ile Phe Asp Asp Val His Gly Ser His Gly
25 30 35

Arg Gln Val Cys Arg Arg His Glu Leu Tyr Val Ser Phe Gln Asp Leu

40	45	50
Gly Trp Leu Asp Trp Val Ile Ala Pro Gln Gly Tyr Ser Ala Tyr Tyr		
55	60	65
Cys Glu Gly Glu Cys Ser Phe Pro Leu Asp Ser Cys Met Asn Ala Thr		
70	75	80
Asn His Ala Ile Leu Gln Ser Leu Val His Leu Met Lys Pro Asn Ala		
90	95	100
Val Pro Lys Ala Cys Cys Ala Pro Thr Lys Leu Ser Ala Thr Ser Val		
105	110	115
Leu Tyr Tyr Asp Ser Ser Asn Asn Val Ile Leu Arg Lys His Arg Asn		
120	125	130
Met Val Val Lys Ala Cys Gly Cys His		
135	140	

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3623 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: pALBP2-781

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2724..3071

(ix) FEATURE:

- (A) NAME/KEY: terminator
- (B) LOCATION: 3150..3218

(ix) FEATURE:

- (A) NAME/KEY: RBS
- (B) LOCATION: 2222..2723

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GACGAAAGGG CCTCGTGATA CGCCTATTT TATAGGTTAA TGTCATGATA ATAATGGTTT	60
CTTAGACGTC AGGTGGCACT TTTCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT	120
TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT	180
AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT ATTCCCTTTT	240
TTGCGGCATT TTGCCTTCCT GTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG	300
CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA	360

TCCCTTGAGAG	TTTCGCC	CCCC	GAAGAACGTT	TTCCAATGAT	GAGCACTTT	AAAGTTCTGC	420
TATGTGGCGC	GGTATTATCC	CGTATTGACG	CCGGGCAAGA	GCAACTCGGT	CGCCGCATAC		480
ACTATTCTCA	GAATGACTTG	GTTGAGTACT	CACCAAGTCAC	AGAAAAGCAT	CTTACGGATG		540
GCATGACAGT	AAGAGAATTA	TGCAGTGCTG	CCATAACCAT	GAGTGATAAC	ACTGCGGCCA		600
ACTTACTTCT	GACAACGATC	GGAGGACCGA	AGGAGCTAAC	CGCTTTTTG	CACAACATGG		660
GGGATCATGT	AACTCGCCTT	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG		720
ACGACCGTGA	CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGAAA	CTATTAACTG		780
GCAGAACTACT	TACTCTAGCT	TCCC GGCAAC	AATTAATAGA	CTGGATGGAG	GCGGATAAAG		840
TTGCAGGACC	ACTTCTGCGC	TCGGCCCTTC	CGGCTGGCTG	GTTTATTGCT	GATAAAATCTG		900
GAGCCGGTGA	GCGTGGGTCT	CGCGGTATCA	TTGCAGCACT	GGGGCCAGAT	GGTAAGCCCT		960
CCCGTATCGT	AGTTATCTAC	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC		1020
AGATCGCTGA	GATAGGTGCC	TCACTGATTA	AGCATTGGTA	ACTGTCAGAC	CAAGTTTACT		1080
CATATATACT	TTAGATTGAT	TTAAAACCTTC	ATTTTTAATT	TAAAAGGATC	TAGGTGAAGA		1140
TCCTTTTG	TAATCTCATG	ACCAAAATCC	CTTAACGTGA	GTTTTCGTT	CACTGAGCGT		1200
CAGACCCGT	AGAAAAGATC	AAAGGATCTT	CTTGAGATCC	TTTTTTCTG	CGCGTAATCT		1260
GCTGCTTGCA	AACAAAAAAA	CCACCGCTAC	CAGCGGTGGT	TTGTTGCCG	GATCAAGAGC		1320
TACCAACTCT	TTTCCGAAG	GTAACTGGCT	TCAGCAGAGC	GCAGATACCA	AATACTGTCC		1380
TTCTAGTGT	GCCGTAGTTA	GGCCACC	TCAAGAACTC	TGTAGCACCG	CCTACATACC		1440
TCGCTCTGCT	AATCCTGTTA	CCAGTGGCTG	CTGCCAGTGG	CGATAAGTCG	TGTCTTACCG		1500
GGTTGGACTC	AAGACGATAG	TTACCGGATA	AGGCGCAGCG	GTCGGGCTGA	ACGGGGGGTT		1560
CGTGCACACA	GCCCAGCTT	GAGCGAACGA	CCTACACCGA	ACTGAGATAC	CTACAGCGTG		1620
AGCATTGAGA	AAGGCCACG	CTTCCGAAG	GGAGAAAGGC	GGACAGGTAT	CCGGTAAGCG		1680
GCAGGGTCGG	AACAGGAGAG	CGCACGAGGG	AGCTTCCAGG	GGGAAACGCC	TGGTATCTTT		1740
ATAGTCCTGT	CGGGTTTCGC	CACCTCTGAC	TTGAGCGTCG	ATTTTTGTGA	TGCTCGTCAG		1800
GGGGCGGAG	CCTATGGAAA	AACGCCAGCA	ACGCCGCCTT	TTTACGGTT	CTGGCCTTT		1860
GCTGGCCTT	TGCTCACATG	TTCTTCTG	CGTTATCCCC	TGATTCTGTG	GATAACCGTA		1920
TTACCGCCTT	TGAGTGAGCT	GATACCGCTC	GCCGCAGCCG	AACGACCGAG	CGCAGCGAGT		1980
CAGTGAGCGA	GGAAGCGGAA	GAGCGCCCAA	TACGCAAACC	GCCTCTCCCC	GCGCGTTGGC		2040
CGATTCA	ATGCAGAATT	GATCTCTCAC	CTACCAAACA	ATGCC	CCCT GCAAAAAATA		2100
AATTCA	AAAAACATAC	AGATAACC	CTGCGGTGAT	AAATTATCTC	TGGCGGTGTT		2160

GACATAAATA	CCACTGGCGG	TGATACTGAG	CACATCAGCA	GGACGCCACTG	ACCACCATGA	2220
AGGTGACGCT	CTTAAAAATT	AAGCCCTGAA	GAAGGGCAGC	ATTCAAAGCA	GAAGGCTTG	2280
GGGTGTGTGA	TACGAAACGA	AGCATTGCC	GTAAGTGCAG	TTCCGGATTA	GCTGCCAATG	2340
TGCCAATCGC	GGGGGGTTTT	CGTTCAGGAC	TACAACGTGCC	ACACACCACC	AAAGCTAACT	2400
GACAGGAGAA	TCCAGATGGA	TGCACAAACA	CGCCGCCGCG	AACGTCGCGC	AGAGAACACAG	2460
GCTCAATGGA	AAGCAGCAAA	TCCCCTGTTG	GTTGGGGTAA	GCGCAAAACC	AGTTCCGAAA	2520
GATTTTTTTA	ACTATAAACG	CTGATGGAAG	CGTTTATGCG	GAAGAGGTAA	AGCCCTTCCC	2580
GAGTAACAAA	AAAACAACAG	CATAAATAAC	CCCGCTCTTA	CACATTCCAG	CCCTGAAAAAA	2640
GGGCATCAAA	TTAAACCACA	CCTATGGTGT	ATGCATTTAT	TTGCATACAT	TCAATCAATT	2700
GTTATCTAAG	GAAATACTTA	CAT ATG CAA GCT AAA CAT AAA CAA CGT AAA				2750
	Met Gln Ala Lys His Lys Gln Arg Lys	1 5				
CGT CTG AAA TCT AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC TTC AGT						2798
Arg Leu Lys Ser Ser Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser						
10 15 20 25						
GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC GCC						2846
Asp Val Gly Trp Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala						
30 35 40						
TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT CTG AAC						2894
Phe Tyr Cys His Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn						
45 50 55						
TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT GTT AAC TCT						2942
Ser Thr Asn His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser						
60 65 70						
AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC AGT GCT ATC TCG						2990
Lys Ile Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser						
75 80 85						
ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA TTA AAG AAC TAT CAG						3038
Met Leu Tyr Leu Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln						
90 95 100 105						
GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC TAGTACAGCA AAATTAAATA						3088
Asp Met Val Val Glu Gly Cys Gly Cys Arg						
110 115						
CATAAAATATA TATATATATA TATATTTAG AAAAAAGAAA AAAATCTAGA GTCGACCTGC						3148
AGTAATCGTA CAGGGTAGTA CAAATAAAAA AGGCACGTCA GATGACGTGC CTTTTTTCTT						
GTGAGCAGTA AGCTTGGCAC TGGCCGTGCT TTTACAACGT CGTGACTGGG AAAACCTGG						3268
CGTTACCCAA CTTAACCGC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA						
AGAGGCCCGC ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCCT						3388

GATGCGGTAT TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCACT	3448
CTCAGTACAA TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC	3508
GCTGACGCGC CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC	3568
GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGA	3623

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 115 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Gln Ala Lys His Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys Lys			
1	5	10	15
Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn Asp Trp			
20	25	30	
Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His Gly Glu Cys			
35	40	45	
Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His Ala Ile Val			
50	55	60	
Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys Ala Cys Cys			
65	70	75	80
Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu Asp Glu Asn			
85	90	95	
Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val Glu Gly Cys			
100	105	110	
Gly Cys Arg			
115			

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 14 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CATGGGCAGC TGAG

14

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GAGGGTTGTG GGTGTCGCTA GTGAGTCGAC TACAGCAAAT T

41

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGATGTGGGT GCCGCTGACT CTAGAGTCGA CGGAATTC

38

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AATTCACCAT GATTCTGGT AACCGAATGC T

31

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GTGGTACTAA GGACCATTGG CTTAC

25

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGACCTGCAG CCATGCATCT GACTGTA

27

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TGCCTGCAGT TTAATATTAG TGGCAGC

27

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGACCTGCAG CCACC

15

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TCGACCCACC ATGCCGGGGC TGGGGCGGAG GGCGCAGTGG CTGTGCTGGT GGTGGGGCT	60
GTGCTGCAGC TGCTGCGGGC C	81

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 73 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGCAGCAGCT GCACAGCAGC CCCCACCACC AGCACAGCCA CTGCGCCCTC CGCCCCAGCC	60
CCGGCATGGT GGG	73

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TCGACTGGTT T	11
--------------	----

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CGAAACCAG

9

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCGACAGGCT CGCCTGCA

18

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCCGAGCGG

10

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CAGGTCGACC CACCATGCAC GTGCGCTCA

29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TCTGTCGACC TCGGAGGAGC TAGTGGC

27